The Way of the Innovator:

Notes Toward a Prehistory of MEDLARS*

BY SCOTT ADAMS

*Biological Sciences Communication Project
The George Washington University
Washington, D.C.*

ABSTRACT

The concept of MEDLARS as a publication/retrieval system was the product of several forces operating during the fifteen years prior to 1961. These included the overriding imperative of medical index publication, the Welch Medical Indexing Research Project, a drive to master the theory of medical subject headings, the innovative impact of coordinate indexing, and the confidence produced by the success of the Index Mechanization Project. By virtue of this experience, MEDLARS became a prime example of a well-understood enterprise seeking a computer application, a circumstance which augured of its success.

GERTRUDE Annan, who honored both us and her distinguished predecessor at the New York Academy of Medicine with the first Janet Doe Lecture in 1967, established the theme which I propose to follow today. Addressing the topic, "The Medical Library Association in Retrospect," she chose to signalize those individuals who singlehandedly pioneered professional advances in medical librarianship. Thus, in our traditions she searched out innovation.

So today, as I respond to your invitation to address "an historical or a philosophical topic," I choose to pursue the same paradoxical course of searching out the seeds of today's innovation in yesterday's history. The complete title of my lecture is: "The Way of the Innovator: Notes Toward a Prehistory of MEDLARS." In it I wish to record my own observations, as a bystander, of the events, forces, benchmarks, and breakthroughs of the fifteen-year period which preceded the award of the MEDLARS contract.

Because I was an observer, and not a participant, I have entitled my lecture "*Notes* Toward a Prehistory." The history itself is still to be written, hopefully by those who participated in it, and if the errors of fact and the misassessments I may make by reason of my secondhand knowledge cry aloud for corrective retribution, or if the National Library of Medicine can be persuaded to update the Rogers-Schullian history, I shall have been amply repaid for my effort.

The late Ralph Esterquest once paid tribute to MEDLARS: "The impact of MEDLARS on the medical library world is not that of the familiar metaphor—the pebble dropped into the pond, casting concentric circles that reach many points on the shore. The impact is no pebble for sure. It is a mighty rock. The waves it will cause will surge and splash for a long time to come. MEDLARS is the great bibliographical breakthrough of our generation" (1). Let me say "Amen!" to that. Not only in the world of medical libraries, but in that of research libraries, of information science, of information systems nationally and internationally, MEDLARS was the rock which raised the level of the ocean.

Whatever may have been, and indeed still are its shortcomings, MEDLARS has grown to be one of the largest openly available information retrieval systems in any field of science in the world. This alone merits an enquiry into its genesis.

The systems conceptualization and basic design of MEDLARS were fundamental to its success. The understanding of the job to be done, in all of its ramifications, culminated in the technical specifications mailed to the systems industry in February 1961, and there my story will end. The completion of the MEDLARS

Bull. Med. Libr. Assoc. 60(4) October 1972

523

contract, the continuing effort to improve *MeSH,* the decentralization, first nationally and later internationally, of the MEDLARS search capability, the conversion to an on-line system and to a national service network: all these highly significant accomplishments I commend to the historical attentions of others.

I am concerned here with an exploration of the infrastructure of MEDLARS, the intellectual and managerial foundations, whose firm placement resulted in the later success. The late Verner Clapp stated that the reason for the success of MEDLARS was that it represented a classic example of a well-defined job seeking a technology, rather than the technology seeking an application (2).

We owe this understanding of the job to be done to a group of pioneers. You will note that I say "group." Dr. Frank Bradway Rogers, who led the group, will be the first to admit that, however pervasive his leadership role may have been, the inputs into the continuing effort which became MEDLARS were many. Indeed, Janet Doe herself contributed to the modernization of the library's indexing program, first as a member of the Survey team assembled by the American Library Association, later as a member of the Committee of Consultants on Indexing, and finally as a member of the Armed Forces Medical Library Advisory Group. The concept of publishing multiple specialized indexes from a common data base, which came to fruition in the MEDLARS recurring bibliographies, can be traced to her suggestion (3).

I have divided my lecture into the following sections:

(1) The Publication Imperative, in which I find the basic motivation which was to drive the National Library of Medicine to MEDLARS;

(2) The Eel of Science, which will concern itself with the pursuit of that slippery and elusive aspect of MEDLARS: the subject indexing of medical papers;

(3) Commotion in the Wings, the offstage background of contemporary noise, without which an understanding of MEDLARS is incomplete; and

(4) The Index Mechanization Project, wherein the publication imperative, the new-found intellectual control, and the contemporary technology were all con-

joined for the first time, leading to MEDLARS.

THE PUBLICATION IMPERATIVE

The history of MEDLARS does not begin, as do so many other aspects of the modern National Library of Medicine, with the 1944 Report of the Survey of the Army Medical Library. The Survey did call for renewed effort to rationalize the functions of the three existing bibliographic services: the *Index-Catalogue,* the *Quarterly Cumulative Index Medicus,* and the *Current List of Medical Literature,* a recommendation which was to be translated into a major program objective of the Army Medical Library. Tacitly, the Survey recognized an assumption of earlier origin: a principal, if not a primary function of the library, was to publish an index to the journal literature of the medical sciences.

The annual reports of Librarians of the Library of the Surgeon General's Office may seem to us now like battle reports from tribal chieftans to the Merovingian kings. A constant theme, however, is their sense of responsibility for continuing the indexing traditions laid down by John Shaw Billings. Indeed, their reports are occupied either with proud statements of *Index-Catalogue* volumes achieved or with apologetic statements for nonpublication. "The clerical force is adequate to keep current work up to date," Col. Percy Ashburn reported successively and without alteration in 1928, 1929, 1930, and 1931, "but there is a great accumulation of back work which has lain dormant for many years" (4).

The indexing must go on, and it must be published. The mission formulated by Billings to acquire, to catalog, and to index the literature of the medical sciences endured long after his departure. When, therefore, the ALA surveyors explored the library, they did not question the library's responsibilities for publishing indexes to the literature of medicine. Index publication must continue, and the library must find a solution to the unification of bibliographic services in medicine (5).

It is significant that one of the first responses of the library's new administration was entitled "The Army Medical Library's Publication Program." In this paper, Dr. Rogers and I traced the long history of indexing publication in the

524

Bull. Med. Libr. Assoc. 60(4) *October 1972*

library and concluded with the announcement that the Committee of Consultants for the Study of Indexes to the Medical Literature had recommended the cessation of publication of the *Index-Catalogue* and the substitution of a current index to current publications (6).

The appointment of this Committee of Consultants was a major event in the library's history. The Survey Committee had recommended that the library establish an advisory committee on the *Index-Catalogue*, with representatives from the American Medical Association, the Medical Library Association, the Bibliographical Society of America, and the National Research Council. In May 1947, Col. Joseph H. McNinch, then Director, proposed the creation of the Committee, and it was established formally on July 7, 1948.

To provide the Committee with an operating arm, which might concurrently investigate problems of priority concern, Col. McNinch arranged a contract through the Army Research and Development Board with the Johns Hopkins University. The contract of Nov. 5, 1948, established the Welch Medical Indexing Project, under the direction of Dr. Sanford V. Larkey one of the first government-sponsored research projects in the field of information science. It investigated, on behalf of the Committee, the size of universe of biomedical literature, and the extent of its coverage by the existing services, the common and disparate features of subject-heading among the services, and the application of machine methods to medical indexing. It is worthy of note that Dr. Larkey hired a young chemist, Eugene Garfield, to head this last-named activity.

The Committee of Consultants for the Study of Indexes to the Medical Literature met many times between its establishment in 1948 and its discharge on May 1, 1952. It addressed a series of questions relating to the publication of medical indexes, including the following:

—What are the indexing requirements of modern medical science?
—Does the *Quarterly Cumulative Index Medicus* meet these requirements? Does any one of the three indexes in existence, or all of them?
—What modifications of existing publications are necessary to meet these requirements?
—What are the characteristics of a good in-

dex, with reference to subject headings, with reference to periodicity, with reference to cumulative issue, with reference to format and printing characteristics?
—What are the possibilities of using mechanical devices to accelerate indexing?

These were both urgent and practical questions relating to the library's traditional responsibility for publishing an index.

For the library in 1945 had taken over the responsibility for the publication of the *Current List of Medical Literature*, which Dr. Atherton Seidell had developed and supported with private funds as a means of announcing the availability of free microfilms. It had done so without adequate calculation of the manpower resources and organization required for its production.

By May 1949, when the Medical Library Association met in Galveston, the weekly issues were a month behind schedule, and the monthly indexes were a year behind. Emergency assignment to the Catalog Division, and a tripling of the staff, were successful in eliminating the backlog, but even as the Committee of Consultants were debating the form of a future index, the fate of the one already undertaken by the library was in doubt. Small wonder that the dominant consideration, becoming at times almost an obsession, was the library's responsibility successfully to publish a current medical index.

Seymour Taine became the Editor of the newly reconstituted *Current List* in August 1949, and the library dedicated itself to a course which was to lead through many vicissitudes to the Index Mechanization Project of 1958–60, and to MEDLARS.

This is not to suggest that the course of the *Current List* was tranquil. Indeed, in 1953 and 1954, Rogers and Taine, with the support of members of the AFML Advisory group, had to fight for its life against conditions and restrictions placed on it by the Bureau of the Budget and higher echelons within the Department of Defense. This dramatic episode, in which for a time all appeared lost, served but to strengthen the library's dedication to index publication.

"I believe," wrote Rogers to members of the Advisory Group, at the time of crisis, "that the *Current List* has the greatest and most immediate social usefulness of all the activities of the library, in proportion to dollar expenditure" (7). The *Current List* survived the crisis. The *Index-*

Catalogue was brought to a successful, albeit truncated, conclusion, and the *QCIM* was terminated. In the final resolution of the duplication of index-publication functions between the library and the American Medical Association, Rogers achieved a cooperative agreement for the latter to publish the annual cumulation of the new *Index Medicus* from copy compiled by the library.

This has been an abbreviated record of the enormous preoccupation of the library with its responsibility for publishing an index to the medical literature of the world. This preoccupation, was a principal force leading to MEDLARS. It accounted, in the Index Mechanization Project, for the priority, assigned to the publication system over the retrieval capability, and in MEDLARS itself for the unit record concept, and for the library's requirement for graphic arts quality computer printout, which led to the development of the Zip Photon. It led also to the exploitation of MEDLARS power to produce specialized recurring bibliographies, and it is still a prime dynamic of MEDLARS.

THE EEL OF SCIENCE

How index learning turns no student pale,
Yet holds the eel of science by the tail.

Alexander Pope

A second theme in the prehistory of MEDLARS is that of the library's long and still-continuing occupation with the theory and practice of subject controls for the literature of biomedicine. Harold Bloomquist has provided us with an excellent review of the attention given to subject cataloging by medical librarians generally (8). Within the National Library of Medicine, there was a special concern for subject controls by virtue of the library's index and catalog publication responsibilities. Anecdotal material, sometimes apocryphal, from the life of John Shaw Billings usually serves to illustrate this traditional concern.

I choose to dip into the library's more remote past for only one passing reference to Billings. Dr. Silas Newcomb, the famous American astronomer, and Dr. Billings were selected to represent the United States at the conference convened by the Royal Society of London in 1896 to plan the *International Catalogue of Scientific Literature*. Otlet and La Fontaine, rep-resenting Belgium at this Conference, vigorously proposed the use of the newly elaborated Universal Decimal Classification as the basis for organizing the *Catalogue,* and Billings just as resolutely opposed them, arguing for indexing under subject rubrics. "But," interposed La Fontaine, "the French do not have a word for 'indexing'. It will be necessary to invent one. Let us call it 'l'indexation' " (9).

Billings prevailed, and the Royal Society developed a classification for the *Catalogue* independent of Dewey and the U.D.C., a circumstance which Bradford, writing nearly fifty years later called "the worst of the calamities" to befall the U.D.C. (10).

At the 1896 conference, Billings offered the following conclusions from a lifetime of medical indexing: "The scientific relations of any subject extend not merely in space of one dimension, but in all the varieties of space of three dimensions, and no possible linear arrangement that can be made will suit the wants of a great majority of enquirers. The scientific investigator is always seeking to make new combinations; he will not be satisfied with any arrangement, nor can any arrangement be made in science, that will be at all permanently satisfactory" (11).

In the years prior to the Survey of 1944, the *Index-Catalogue* had become the bibliographic tail which wagged the library dog. The surveyors were hypercritical both of the internal inconsistencies in the organization of the *Index-Catalogue,* and of its indifference to a standard list of subject headings for medicine. In what is perhaps one of the more simplistic of statements made about a complex intellectual function, the surveyors called for such a standard list, and proposed that "headings chosen should be those most commonly known and used in this country, and adequate cross references should be provided to round out the system" (5). They proposed further that "the list of subject headings that has been worked out by the *Quarterly Cumulative Index Medicus"* be taken as the basis of the new list. Parenthetically, it might be noted that the *QCIM* Authority List, dating to 1935, is one of the earliest examples of a thesaurus for an indexing system.

For a library which had never had a book classification, the Surveyors recommended the development of a modern scheme "to be based (hopefully) on a thorough revision of the Li-

brary of Congress schedules for medicine, with certain basic ideas drawn from the Cunningham classification" (5). Mary Louise Marshall of the Survey team was given a contract to develop this new classification.

A preliminary version of the new classification was delivered to the library in 1948 and was promptly referred by the library's Director, Col. Joseph H. McNinch, to Dr. (then Major) Rogers, working on his master's degree in library science at Columbia University. Dr. Rogers's introduction to the subject control of library materials was thus far from academic; with his extraordinary ability to organize masses of detail, he inundated the library with a spate of revisions, which were incorporated in time for a first edition of the *Classification* in September 1949.

In the meantime, even prior to Dr. Rogers's involvement with the *Classification,* the library had taken a preliminary step to implement one of the Survey's recommendations. On December 12–13, 1947, the library conducted a Symposium on Medical Subject Headings for the dual purpose of addressing the questions: What principles may be found for developing subject heading terminology in the medical sciences? and What are the recognizable differences in principles to be applied to the subject heading of monographic and periodical literature?

Janet Doe presented a paper "A Critical Review of Existing Medical Subject Heading Lists," and Sanford Larkey, soon to contract with the Army Research and Development Board for the Welch Medical Indexing Project, a paper on "Introduction to the Problems of Medical Subject Heading." An extracurricular activity of this symposium, however, is particularly prophetic. The minutes of the Symposium contain this note: "During the recess before discussion, Col. McNinch led a group through the Medical Statistics Division, Office of the Surgeon General, where the IBM punch card machinery was in operation. Possible application of punch card controls for medical bibliography was suggested" (12).

The 1947 symposium came at the low point in the library's indexing fortunes. The backlog problem had in fact become so acute that the MLA meeting in Galveston in April 1948 took formal action to urge the library to catch up. Volunteers were welcomed at that time, with

Helen Bayne preparing the May 1948 and Mary Louise Marshall the June 1948 subject indexes to the *Current List.*

By heroic efforts, the indexing backlogs were reduced by the time Dr. Rogers became Director of the library on October 21, 1949. Seymour Taine had been made Acting Editor on August 19 of that year, and the staff of the *Current List* had been considerably augmented. With the major work of the revision of the classification behind him, Dr. Rogers could devote time to the stubborn intellectual problems of subject headings.

The Welch Medical Library Indexing Project had as one of its objectives exploration of "The theory and practice of subject heading (nomenclature) and classification (coding) as they concern medical literature." Helen Field, Williamina Himwich, and Sanford Larkey, working on the Project, had explored the principles underlying the subject heading practices of the *Index-Catalogue* and the *Quarterly Cumulative Index Medicus.* They found the categorization of the terms used convenient for purposes of analysis, and created sixteen categories, a prototype of the *MeSH* categories (13).

From the outset, Dr. Rogers was confronted with the necessity of solving problems of subject heading in two separate, but related contexts: that of providing a subject catalog for the library and that of creating a subject indexing authority list for the *Current List.* The problem had many interesting ramifications. Not only were there differences of traditions of library cataloging and scientific indexing; there were also fundamental differences in user groups and their purposes, and in the characteristics and forms of the literatures.

The revitalized *Current List* had depended on the *QCIM* list for its subject heading authority. This list needed updating and was not suited to the new indexing style of the library's "Index." It became imperative to implement the Surveyors' recommendation and to develop an authority list proper to the *Current List.*

The theoretical work done on subject headings in the Welch project was discontinued, and a team including Dr. Rogers, Mr. Taine, Mrs. Gertrude Butler, and Miss Helen Field undertook a crash effort to produce on 6335 IBM cards, by October 1951, the list which became the first *Subject Heading Authority List*

(*SHAL*). The Field categories again proved useful in dividing up the work.

With this introduction, both practical and urgent, Dr. Rogers set for himself an intellectual challenge which has occupied his professional career—to master the principles underlying the construction and use of formal systems of subject headings in the medical sciences. I need only remind you that the subject of his own Janet Doe Lecture in 1968 had to do with the impact of changes in *MeSH* on the subject catalog of his library.

Dr. Rogers's first conclusions were published in a collection edited by his Columbia mentor and colleague, Dr. Maurice Tauber, *The Subject Analysis of Library Materials* (14).

Noting that subject control of the periodical literature is a preeminent requirement in science, Dr. Rogers proceeded to analyze the likenesses and differences between subject headings for periodical indexes and book catalogs. He concluded that the similarities were basic (a conclusion which was to lead later to the use of *MeSH* for both the *Index Medicus* and the *NLM Catalog*), and that the principal differences arose from the need for topical subdivision in indexes, as well as a more intensive degree of subject analysis. This theme and its variations were to preoccupy Dr. Rogers until he resolved the conflict by creating *MeSH* (or *MSH*, as it was initially known) as a combined indexing and subject heading authority list in 1959.

In the meantime, a new and a very powerful influence on the library emerged. Dr. Mortimer Taube, after serving as Chief of the Science and Technology Division, Library of Congress, and Deputy Chief for Scientific Information of the Atomic Energy Commission, had established Documentation, Inc., as a vehicle for expounding, marketing, and further developing his revolutionary concept of coordinate indexing. Many of us on the Washington scene were fascinated with the simplicity of Taube's concept of posting items on subject terms, instead of subject terms on items, as we had been taught. We watched with great interest his installation of Uniterm indexing at the Armed Services Technical Information Agency (ASTIA) and read avidly his *Studies in Coordinate Indexing*.

Dr. Rogers became one of the staunchest interpreters and defenders of Taube in the library profession. His two reviews of Taube's work showed him not only to have mastered the implications of Taube's original thought but also to be an active champion for Taube, seeking to stir up the library community to this powerful new idea. Rogers did not follow Taube blindly, however. As friends who had taken each other's measure, they must have argued long over the principles underlying Uniterms. "What was and is essentially wrong with the Uniterm system," wrote Rogers in his staunchest exposition of Taube's work, "is not its unusual posting system, and certainly not its central coordinate indexing concept, but its employment of article-derived catchwords in preference to a carefully chosen controlled vocabulary of terms" (15).

Rogers's philosophy of medical subject headings, which he had worked out through wrestling with might and main with this most challenging and elusive of topics, is best expressed in the prefatory matter to the first edition of *MeSH*, 1960. Subject cataloging and periodical indexing are, in their major dimensions, identical processes, and a single authority should be used for both. Headings should be considered as pointers, not as descriptive labels. Topical subheadings, as substitutes for phrase headings, can be standardized, and should be, in contrast to specific main headings that generally represent broad concepts.

It is interesting to note, in the 1960 Preface, the influence of the coordinate indexing concept: Subject headings, being pointers, "the intersect of two or more such pointers will define a very specific subject" (16).

The second edition of *MeSH* was prepared specifically in anticipation of its use in the search operations of MEDLARS. Winifred Sewell joined the library in May 1961 as the first full-time editor. The sixty-seven subheadings of *MeSH* 1, in combination with the main headings, exceeded the estimated capability of the developing system, and after test with a more finite set of subheadings had failed, it was decided to abandon topical subheadings and to substitute where practicable precoordinated terms. The categories were thoroughly reviewed, and authoritative nomenclatures checked for the most generally accepted terminology.

Over the years immediately preceding MEDLARS, much effort went into the preparation of the *MeSH* vocabulary for machine retrieval. *MeSH* was evolving into a controlled list of

descriptors, and away from a list of conventional subject headings. "The new subject heading," said Rogers in reporting on computer developments to the Board of Regents, "says in effect, 'There it is, right there'. Two or more of the new subject headings will indicate a precise point more satisfactorily, as by an intersection of vectors. What is more, they lend themselves more readily to adequate control, and maintenance of consistency in the retrieval vocabulary is a necessity" (17).

Thus did the library prepare itself intellectually over the years, through trial and error, theory and practice, for the advent of MEDLARS. Since Billings, the vocabulary had been built laboriously from the living literature; the structure had been evolved through both trial and error and theory.

Despite this long record of experience, the testing of the vocabulary under operating conditions showed need in many areas for improvement. The need continues, as evaluation and testing have shown. But the main outlines and structure of *MeSH* have survived, a credit to the effort which Dr. Rogers expended on this most intractable of problems.

COMMOTION IN THE WINGS

This is not the place, nor do I have the time, to describe at any length the onrush of contemporary events in the world of information retrieval, yet no discussion of the prehistory of MEDLARS can be complete without an understanding of the frenetic environment in which the library's concepts matured. Memex, Vannevar Bush's electronic adaptation of the World Brain proposed by H. G. Wells, had captured the imagination of all. With the creation of ENIAC and SEAC, the electronic revolution was upon us, and the computer age had been born. Imaginations were unbridled; potential applications to bibliographic management proliferated; initiatives abounded, and research and development funds flowed freely.

Rapid and spectacular as was the growth of data processing capability and systems technology during these years, the growth of projects conceived to apply them to the "information problem" was even more dramatic. Industry aided and abetted academic enthusiasts; information processing technology was overadvertised and oversold on every hand. At dozens of

meetings and symposia, and at industrial tutorial retreats, statements descriptive of potential applications of mechanized information handling were presented in the declarative mood, rather than in the more correct subjunctive. These were indeed exciting, if frustrating, times.

One can even be nostalgic about some of these developments. Where are the systems of yesteryear? A review of the National Science Foundation's publication *Current Research and Development in Scientific Documentation* for these early years reveals dozens of names, once billed as universal solutions, now fading into history. Documentation, Inc., produced COMAC, Kodak developed MINICARD, while Magnavox worked on MAGNACARD. The CIA engaged in Project Walnut. The National Bureau of Standards perfected a "peek-a-boo" system for controlling the literature of basic instrumentation. The National Research Council organized the Chemical Biological Coordination Center for punched card control of information on the biological action of chemical compounds. Western Reserve promoted "semantic factoring" as the ultimate intellectual control over information and constructed the WRU Searching Selector to use it. The Zator Company developed Zatocoding, while George Washington University worked out Tabledex.

There were claims and counterclaims, rivalries and polemics, and one scarcely knew whether to place bets on a sure thing, or whether to play the field.

The first Sputnik added greatly to the confusion. Politicians rapidly convinced themselves, not without help from American information specialists, that the spectacular Soviet success was due to a superior scientific and technical information service, and charted a course of national competition to catch up with, and then best, the Soviets. The management of scientific and technical information programs by federal agencies came under the critical scrutiny of both houses of Congress; oil was poured on the fires of competition, and technological confusion compounded with political commotion.

The National Library of Medicine inevitably drew the attention of Senators and Congressmen who, on more than one occasion, expressed their impatience with the library's inability to match some of the more imaginative solutions to the science information problem then being pro-

Bull. Med. Libr. Assoc. 60(4) October 1972

529

posed by the library's contemporaries in academia and industry.

But during these years the overbearing problems of the library were not technical ones relating to information retrieval. As a responsible operating agency with publication deadlines to meet, the library's problems were concerned with getting the work done and on time. They were major organizational and management ones: the interagency struggles, as well as those internal to the Department of Defense, for the stewardship of the library; resistance to the transfer of the library to Chicago, the drafting and passage of legislation establishing the National Library of Medicine; and, above all, that supreme imperative, the new building.

Inevitably, Dr. Rogers and his associates were involved in the new information technology as critics and advisors. They attended the documentation and information science meetings and symposia and served as consultants on such projects as the Cardiovascular Literature Indexing Project of the National Research Council. They were in the world of systems, and yet not quite part of it. As a responsible and overburdened manager, Rogers was appalled by the nonsense and waste which he identified in so many of these efforts to reinvent wheels, and made no bones about saying so. This experience not only kept him and the library currently aware of the developing technology; it sharpened his judgment and firmed his convictions on what would or would not work.

THE INDEX MECHANIZATION PROJECT

The new *Current List of Medical Literature* had begun its publication under Seymour Taine's editorship in 1950 in three separate sections: the article registry, and the subject and author indexes. A serial numbering device was used to link the two indexes to the full citation found in the registry. In nine years of successful production, the index had seen a steady growth; journals covered had increased from 1,225 to 1,600, and articles indexed reached 110,000.

But a 1957 survey by Estelle Brodman and Seymour Taine for the International Conference on Scientific Information had estimated the current world production of medical papers at 220,000, and the NLM was under compulsion to increase its coverage. However, the larger the issues became, the more cumbersome the proc-

ess of referral and rereferral between the index sections and the article registry. Clearly a new format was essential.

This practical problem led to a decision to undertake a development project to apply the best of the current information processing technology to the production of the *Current List*. An application for $73,800 to support a two-year developmental project was submitted to and approved by the Council on Library Resources in the Spring of 1958 (18).

Index publication was still the imperative. The project's objectives were those of "developing and demonstrating in the field of medicine improved methods for the rapid and efficient publication of comprehensive indexes to the literature of broad scientific fields, with simultaneous provision for meeting the requirements of specialties within these fields, making use of hitherto unutilized mechanical applications" (19).

It might be noted that the secondary objective, the "simultaneous provision" was conceived of in terms of "encoding unit-citation cards . . . to make them immediately available for whatever limited indexes that might be desired (for example: cancer chemotherapy listing; cardiovascular drugs listing)." The objective was closer to what Janet Doe had proposed in 1952 and to the recurring bibliographies of MEDLARS than it was to bibliographic search on demand. In passing, it may be noted that an Advisory Committee to the project included both Dr. Eugene Garfield and Dr. Sanford Larkey, a provision for continuity between the Welch Project and the library's new venture.

The published report of the Index Mechanization Project is generally conceded to be one of the more candid records of the successes and failures of an information systems development project ever written. The operational history, demonstrating the almost daily operation of Murphy's Law (if anything can go wrong, it will) has been the envy of many report writers.

The project accomplished its primary objective brilliantly, and the published *Index Medicus* from 1960 through 1963 is its monument. It failed in its efforts to achieve the type of subject retrieval that would permit the publication of specialized bibliographies.

But, as Dr. Rogers stated in his *Annual Report* for 1960, "As for the ill-fated bibliographic

530

Bull. Med. Libr. Assoc. 60(4) October 1972

retrieval system—despite the negative outcome, much valuable experience and knowledge has accrued which may well be of greater significance in the long run than the successful publication program. . . . NLM received the impetus for further investigation of the application of more sophisticated mechanized techniques, which will continue to influence the library in the future" (20).

I believe this conservatively worded forecast of things to come owes its origin to an educational event which occurred in the fall of 1958. The Central Intelligence Agency had sent Joseph Becker to UCLA to study information retrieval and to familiarize himself with the advances made by the Southern California computer industry. Foster Mohrhardt, then with the National Agricultural Library, at Becker's suggestion, persuaded Dr. Rogers and Verner W. Clapp to pay Becker a visit. The four friends visited many California firms, discussed information retrieval systems, and argued over a ten-day period.

Dr. Rogers had gone to California a skeptic, believing without doubt that the California systems approaches were no better and no worse than those he had criticized on the East Coast. He returned a changed man, with a new set of convictions. As in the case of subject headings, he had wrestled the problem to the mat and won. He had met the computer, and had become its master. Indeed, so thoroughly had he mastered the principles of the computer that he presented to the NLM Board of Regents at the instigation of Dr. Michael DeBakey, then Chairman, a tutorial on the "Consideration of the Use of Computers in Bibliographical Tasks" (21).

I can hear him now as he elucidated programming to the Board members. "If we want to tell a clerk to put a letter on the table, we can say, 'Put this over there'. But suppose we have to instruct a mechanical robot to do the same thing. We would give him a long sequence of minute instructions: 'Extend right hand. Open fingers. Accept letter. Close fingers. Lower arm. Turn 180 degrees. Place left foot ahead of right foot. Place right foot ahead of left foot. Etc. Etc. Halt. Extend right arm. Open fingers. Turn. Etc.' "

This was what the computer business all boiled down to: detailed instructions to accomplish a predetermined objective. And meticulous attention to detail and to accuracy is present both in library cataloging and in computer programming. Medical indexing and the computer obviously were made for each other.

This new-found confidence in the library's ability to understand and utilize the computer was added to the earlier foundations: a long and successful record in the management of indexing and of index publication, a hard-earned mastery over the principles of subject heading, an enthusiasm for the philosophy of coordinate indexing, and an outstanding success in the Index Mechanization Project. All the preconditions necessary to MEDLARS were here but one, the necessary financial resources.

These came from a quite unexpected source. The law establishing the National Heart Institute within the National Institutes of Health authorized the Surgeon General to "establish an information center on research, prevention and treatment of heart disease." In an effort to secure guidance on what he might do most effectively to implement this statutory responsibility, Dr. James Watt, Director of the National Heart Institute, in January 1960 commissioned a study. Mr. J. Douglas Knox, the contractor-investigator, noting the scope of the library's coverage of the biomedical literature, the success of its Index Mechanization Project, and its interest in further developmental work, concluded that "establishing an effective flow of scientific communications is a problem of documentation which logically suggests a collaborative effort with the NLM." He recommended that the National Heart Institute establish a joint working group with NLM to explore advanced techniques for information retrieval (22).

The Knox Report is amazingly prophetic. Knox proposed "a collaborative plan with the National Library of Medicine, whereby NLM would be responsible for the mechanized storage of the world's biomedical literature, and selective retrieval geared to the needs of the cardiovascular scientist; and whereby the National Heart Institute would be responsible for planning utilization programs for the basic bibliographical and documentary data produced automatically by NLM, with servicing of requests from regional centers by an electronic network joined with the master center at Bethesda, from which current lists, abstracts and facsimile documents could be rapidly disseminated."

The Knox Report led to negotiations between the library and the National Heart Institute, and at its fall meeting in 1960 the National Advisory Heart Council approved the transfer of $500,000 from NHI to NLM for the purpose of initiating the development project.

The fall months of 1960 were spent in drafting the technical specifications for MEDLARS. The necessary arrangements were made with the Public Health Service contracting officials for advertising for bids.

It has been stated that in the Index Mechanization Project the library had as its objective the development of a publication system from which a retrieval subsystem might be derived. The serial search of the million punched cards produced annually, however, made the subsystem impractical.

In contrast, MEDLARS was conceived of as a system based on individual machine-readable citation records from which both a publication system and a retrieval system might be elaborated with equal facility. This is evidenced in the primary objectives associated with the statement of specifications, which included among the desiderata not only an increase in the volume and speed of indexing publication, but also "the prompt and efficient servicing of relatively complex demand requests for bibliographic information, and recurring bibliographies on such topics as 'diseases of the cardiovascular systems.'"

On Feb. 2, 1961, the technical specifications, accompanied by invitations to submit proposals, were sent to forty selected commercial and nonprofit organizations. The word spread throughout industry, and before the deadline for submission of bids arrived, some seventy-two companies had indicated their interest. Twenty-five proposals in all were received by the deadline, April 24, 1961 (23). It is worthy of note that the industry was highly complimentary over the thoroughness, the precision, and the clarity with which the specifications were drafted, a circumstance which paid tribute to the library's comprehension in detail of the job to be done, and of the ability of the computer to perform it.

It has been fashionable to comment that while John Shaw Billings conceived the *Index Catalogue* in 1876 and the punched card in 1883, it took over eighty years before the two were successfully married. For those who may believe that successful information retrieval systems are products of instant parthenogenesis, I wish it noted that even after the punched card technology and medical indexing were introduced to each other in the Welch Medical Indexing Project, it took ten years of ardent and arduous courtship before the union could occur.

REFERENCES

1. ESTERQUEST, RALPH T. Discussion of paper presented by Taine, Seymour I. Bibliographic data processing at the National Library of Medicine. In: Clinic on Library Applications of Data Processing, University of Illinois. Proceedings of the 1963 Clinic. Champaign, Ill., The Illini Union Bookstore, 1964. p. 124–125.
2. CLAPP, VERNER W. Personal communication.
3. DOE, JANET. Methods for medical indexing. Bull. Med. Libr. Assoc. 39: 23–27, Jan. 1951.
4. U. S. ARMY. Surgeon General's Office. Annual reports, 1928, 1929, 1930, 1931.
5. METCALF, KEYES D. [and others]. The National Medical Library; Report of a Survey of the Army Medical Library. Chicago, American Library Association, 1944.
6. ROGERS, FRANK B., AND ADAMS, SCOTT. The Army Medical Library's publication program. Texas Rep. Biol. Med. 8: 271–300, Summer 1950.
7. ROGERS, FRANK B. Memorandum to Members of the AFML Advisory Group, 5 January 1954.
8. BLOOMQUIST, HAROLD. Cataloging and classification of medical library materials, 1946–1956. Ten years of progress and problems. Bull. Med. Libr. Assoc. 47: 28–47, Jan. 1959.
9. INTERNATIONAL CONFERENCE ON A CATALOGUE OF SCIENTIFIC LITERATURE. Report of the Proceedings . . . July 14–17, 1896. London, The Royal Society, 1896, p. 27.
10. BRADFORD, S. C. Documentation. London, Crosby, Lockwood & Son, 1953, p. 140.
11. *Op. cit.*, p. 78–79.
12. Symposium on Medical Subject Headings. Minutes of the Symposium . . . Dec. 12–13, 1947. Washington, Army Medical Library, 1947, p. 3.
13. FIELD, HELEN G. [and others]. Final Report on Subject Headings and on Subject Indexing. Baltimore, Welch Medical Library Indexing Project, The Johns Hopkins University, 1955.
14. ROGERS, FRANK B. Application and limitations of subject headings: the pure and applied sciences. In: Tauber, Maurice F., ed. The Subject Analysis of Library Materials. N.Y., Columbia University School of Library Service, 1953.
15. ROGERS, FRANK B. Review of Taube, Mortimer. Emerging solutions for mechanizing the storage and retrieval of information. (Stud-

ies in Coordinate Indexing, vol. 5) Coll. and Res. Libr. 21: 489–492, Nov. 1960. See also, Rogers, Frank B. Review of Taube, Mortimer. Studies in coordinate indexing. Bull. Med. Libr. Assoc. 42: 380–384, July 1954.

16. NATIONAL LIBRARY OF MEDICINE. Medical Subject Headings. First Edition. Washington, Government Printing Office, 1960. p. i–xix.

17. ROGERS, FRANK B. Consideration of the Use of Computers in Bibliographical Tasks. Unpublished abridged version of a paper presented at the Meeting of the National Library of Medicine Board of Regents, April 11, 1960.

18. TAINE, SEYMOUR I. New program for indexing at the National Library of Medicine. Bull. Med. Libr. Assoc. 47: 117–123, Apr. 1959.

19. NATIONAL LIBRARY OF MEDICINE. Index Mechanization Project, July 1, 1958–June 30, 1960. Bull. Med. Libr. Assoc. 49, no. 1, pt. 2. Jan. 1961.

20. NATIONAL LIBRARY OF MEDICINE. Annual Report for the Fiscal Year 1960. Washington, National Library of Medicine, 1960. p. 41.

21. ROGERS, op. cit., p. 6–7.

22. KNOX, J. DOUGLAS. Report of a Study on Collection and Dissemination of Heart Research Information. Bethesda, National Heart Institute, 1960.

23. NATIONAL LIBRARY OF MEDICINE. Annual Report for the Fiscal Year 1961. Washington, National Library of Medicine, 1961. p. 27–29.